

living planet symposium | BONN

23–27 May
2022

TAKING THE PULSE
OF OUR PLANET FROM SPACE

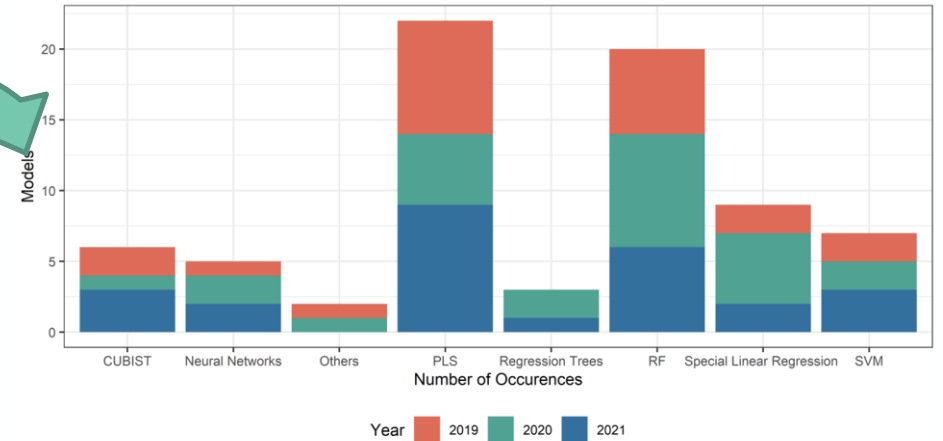
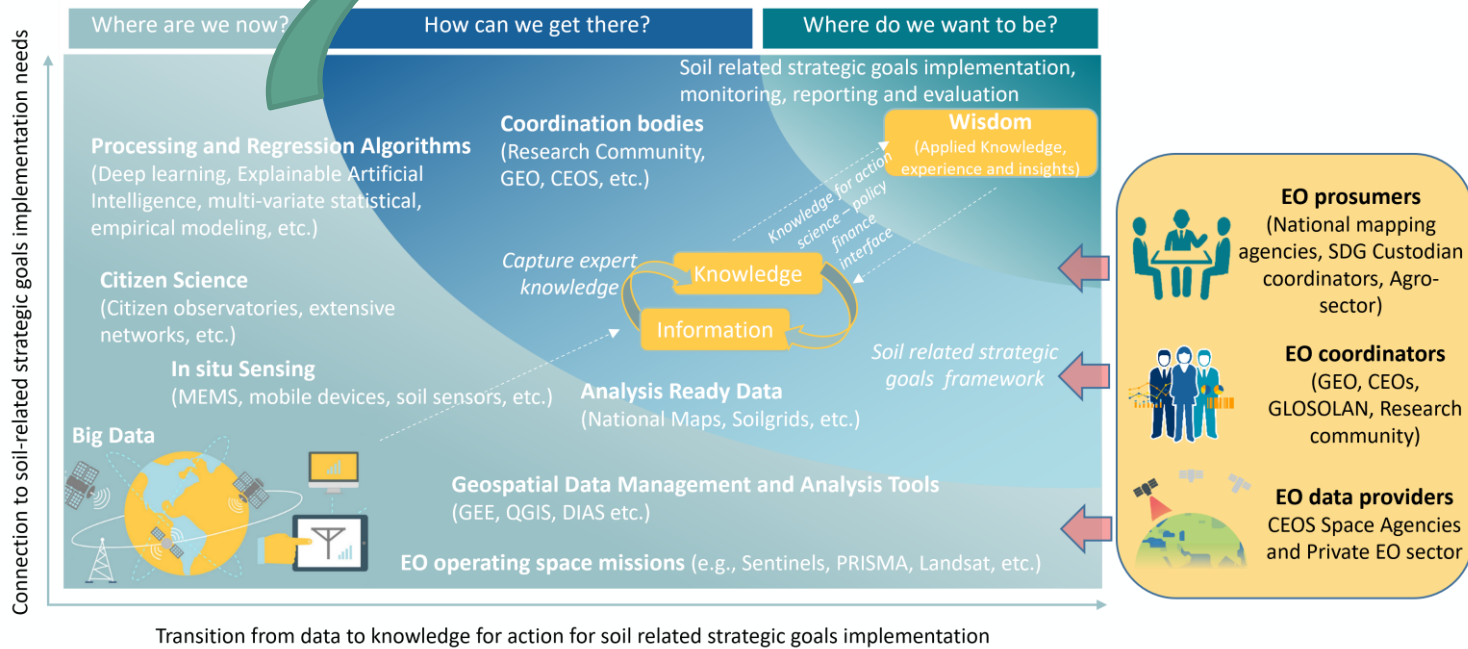


Convolutional neural networks for soil organic carbon mapping from Sentinel-2 satellite imagery; a case study in Bavaria state

Nikolaos Tziolas, Dr. Uta Heiden, Klara Dvorakova, Dr. Pablo d'Angelo, Simone Zepp, Prof. Dr. Bas van Wesemael

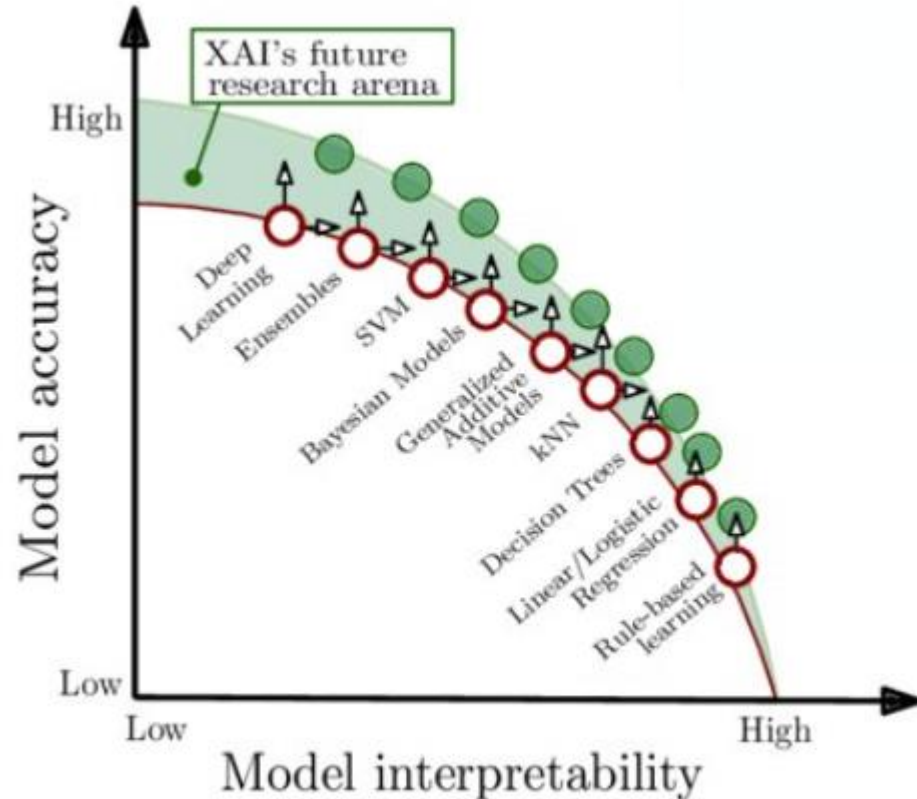
27/5/2022

<https://doi.org/10.3390/rs13214439>



- PLS and Random Forest are certainly the most popular ML algorithms;
- In 2015 has been applied for first time DL architectures for soil property prediction;
- There is an increase in the use of Neural Networks (e.g., DL)

Pillars that should drive the DL architecture development

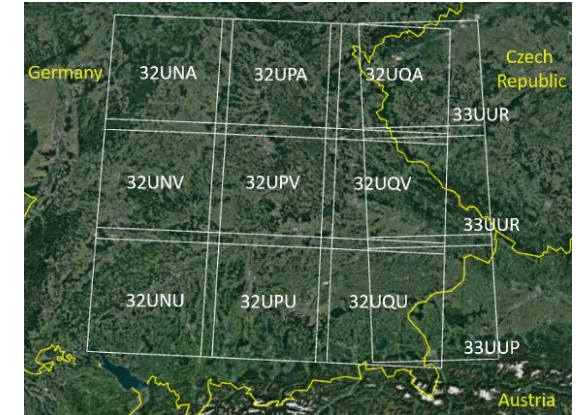
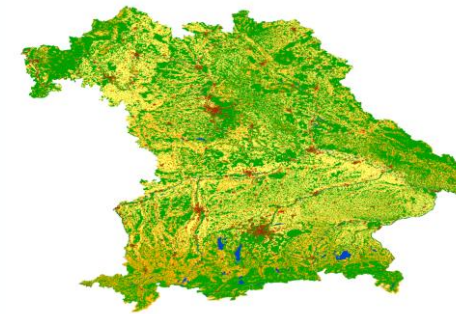
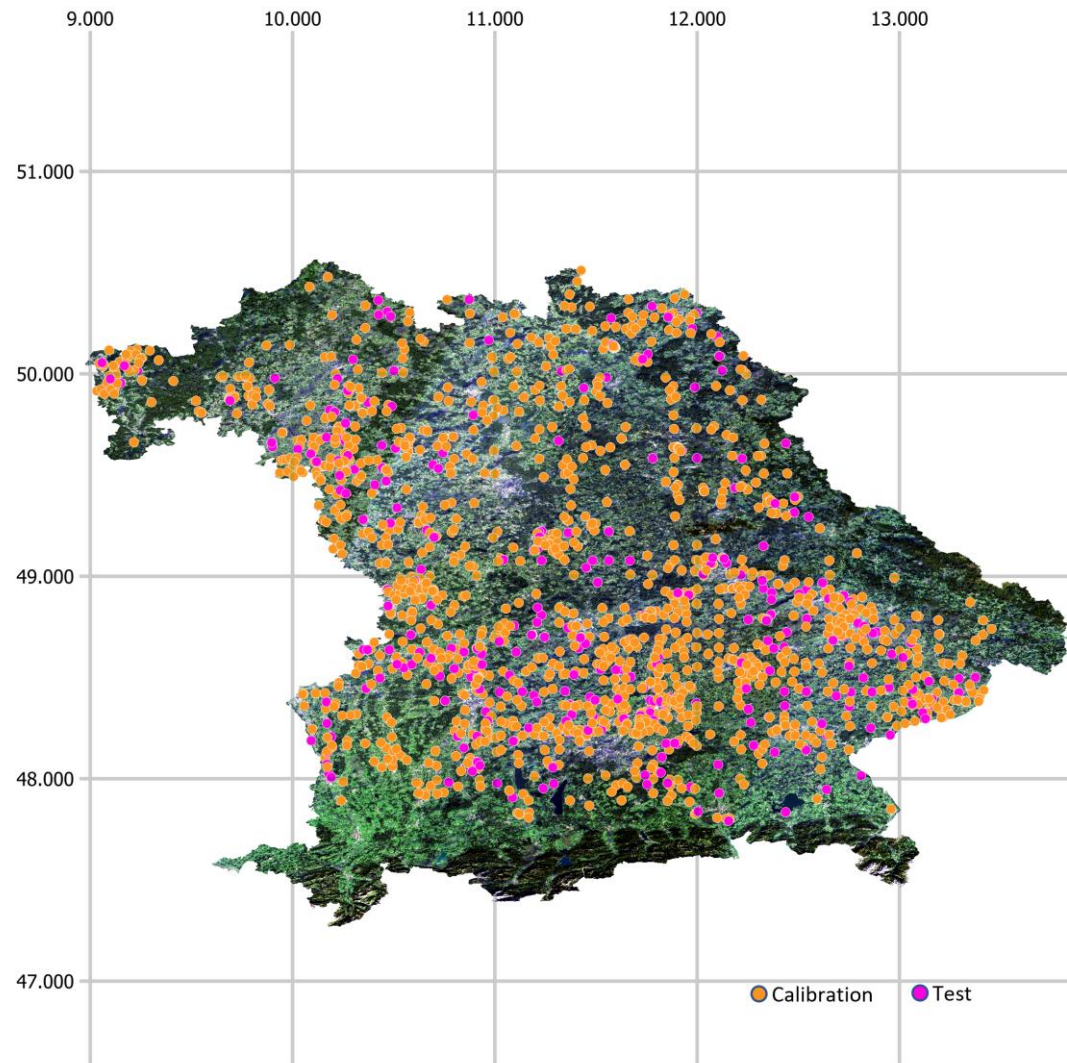


- DL is not a **panacea**
- **Move from black to glass box models (XAI);** Explainability is important for debugging AI models and making informed decisions (<20% present explainability or mentioned it's importance)

Interpretability: Post-hoc explainability techniques; Interpretability-driven model designs

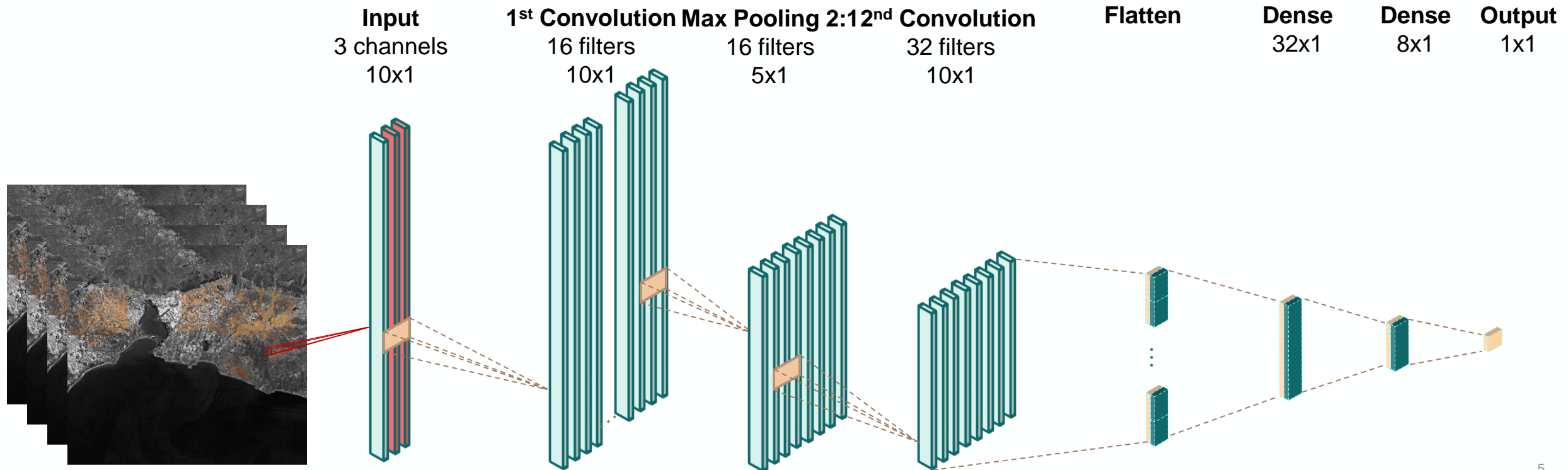
Accuracy: Hybrid modelling approaches; New explainability-preserving modelling approaches

(dx.doi.org/10.1016/j.inffus.2019.12.012)

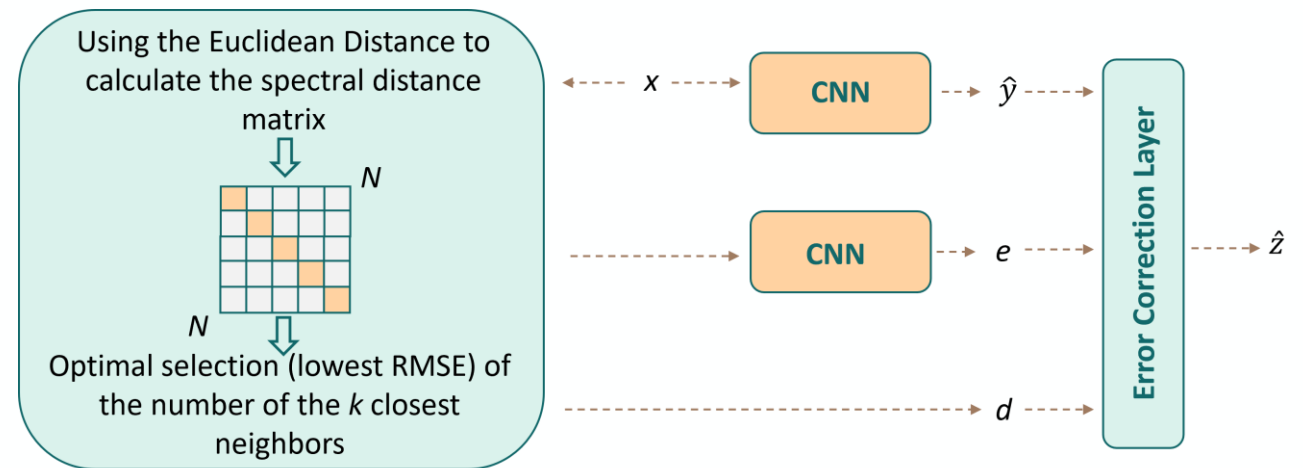
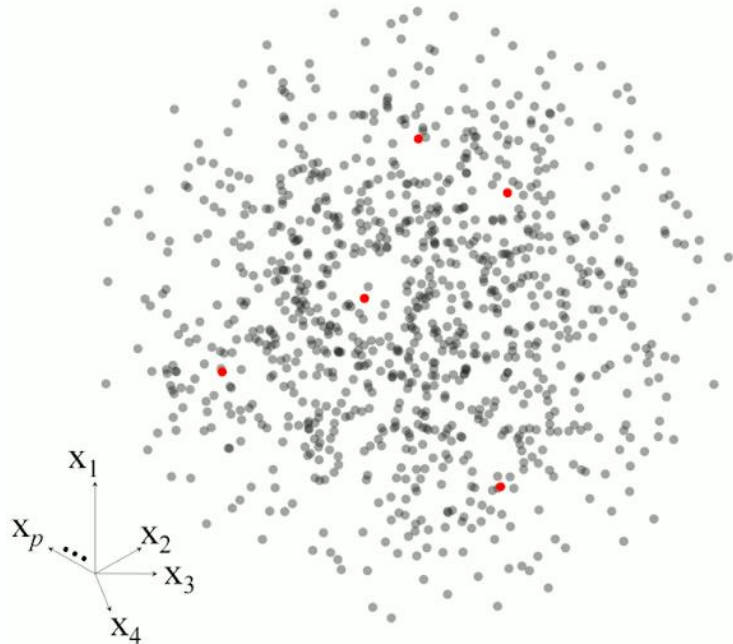


- **1933 topsoil samples** (LUCAS, Bavarian Environmental Agency and State Research Center for Agriculture); SOC content ranges from 0.26 to 180 g · kg⁻¹
- **Sentinel-2** data (exc. 60m bands) from 12 tiles; cloud coverage <80%; range from 2018-2020; i) spring; ii) spring-autumn; and iii) full months
- Processing with **Soil Composite Mapping Processor** by DLR

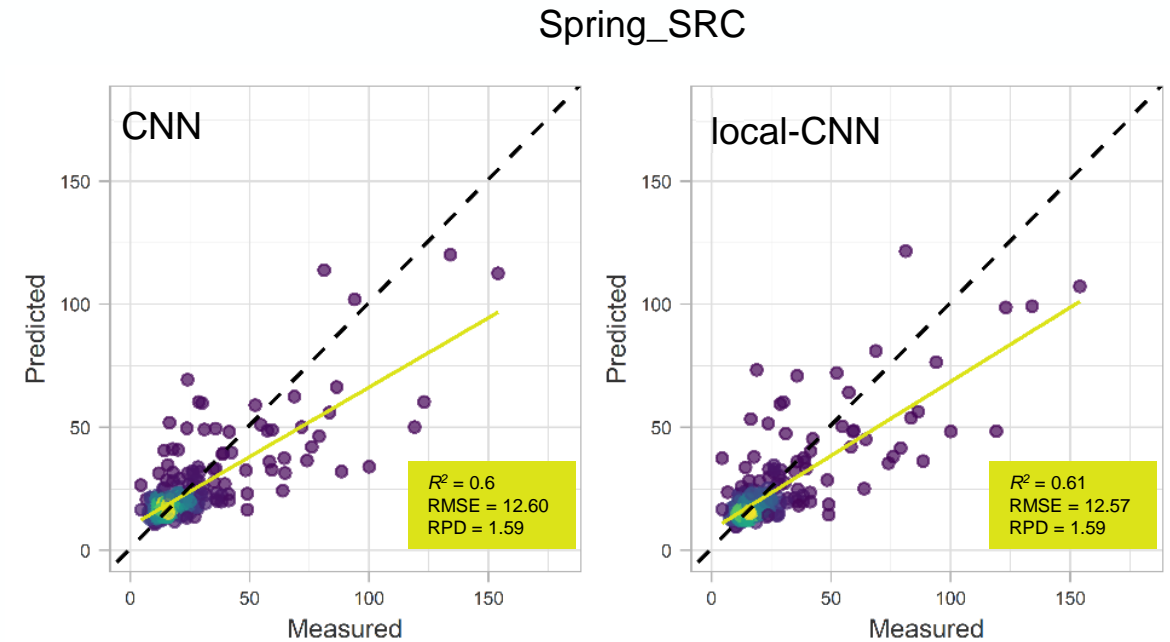
- **Shallow deep learning** architecture able to handle multispectral data, supporting also multi-output predictions;
- Exploits the **complementary information** contained from multiple spectral sources (no need to find the best pre-treatment);
- Address the issue of **interpretability**;
- Evaluate the inclusion of spectral indices (e.g., NDVI, NBR2 etc.) as additional predictors (**CNN with indices**)



- **Local error correction mechanism**, where information from a global model is used to localize multiple models (utilize the $k \in [10, 200]$ nearest neighbors)
- Spectral distance calculation and closest neighbor selection by Euclidean Distance

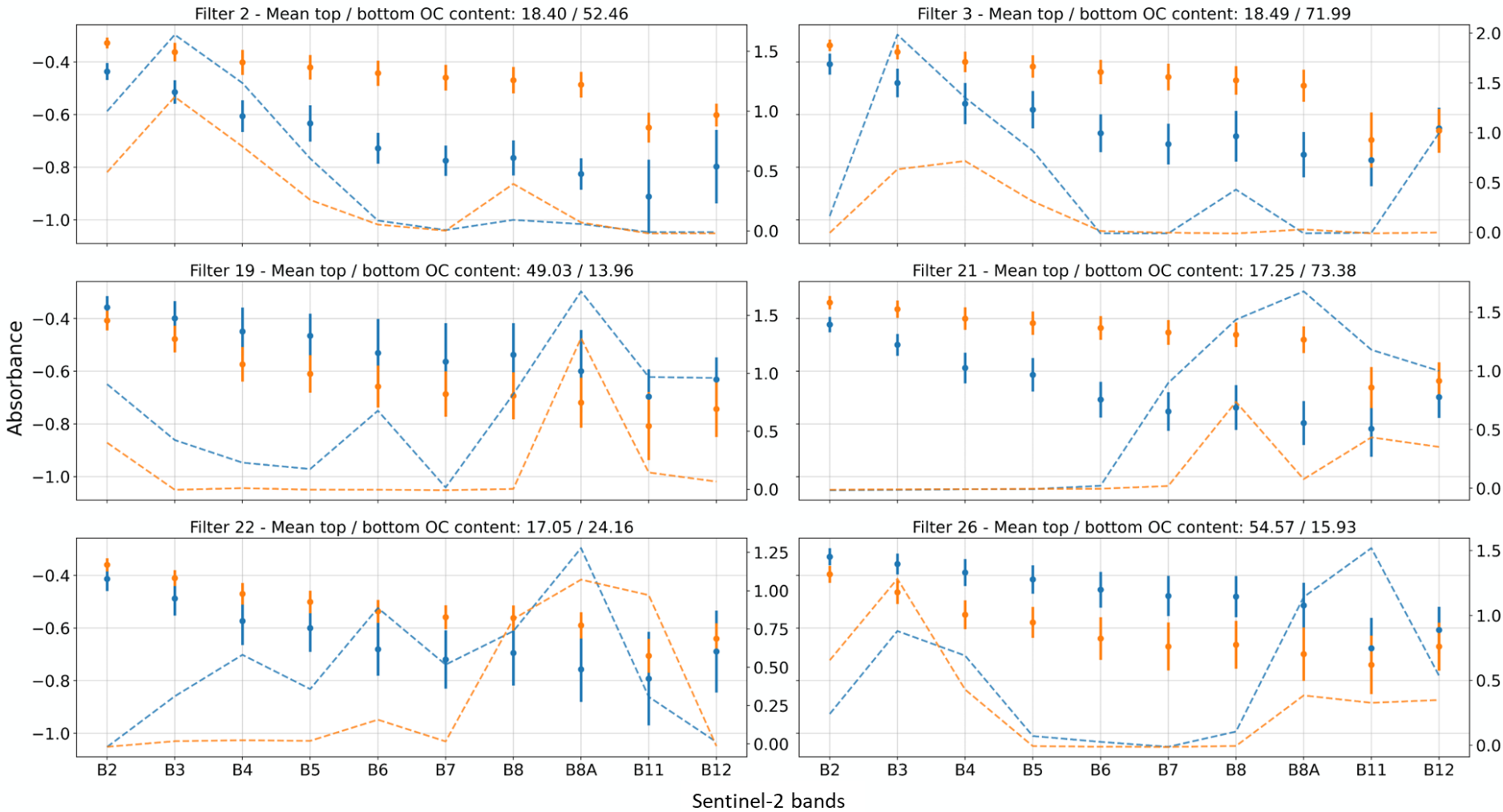


| Composite | CNN | | | CNN with indices | | |
|------------|--------------|----------------|-------------|------------------|----------------|------|
| | RMSE | R ² | RPD | RMSE | R ² | RPD |
| Full_SRC | 12.65 | 0.60 | 1.58 | 12.93 | 0.58 | 1.52 |
| Autumn_SRC | 12.84 | 0.59 | 1.56 | 13.27 | 0.59 | 1.51 |
| Spring_SRC | 12.60 | 0.60 | 1.59 | 13.55 | 0.54 | 1.48 |



- Best performance for the Spring soil reflectance composites (SRC)
- There is no need to include vegetation indices as additional spectral features in CNN;
- Statistical **marginal improvement (~1%)** due to the proposed **error-correction scheme**;
- The herein proposed **CNN scored significantly better** compared to current SOTA models for Sentinel-2 data (PLS: $R^2=0.49$, RMSE = $13.76 \text{ g} \cdot \text{kg}^{-1}$, RPD = 1.4).

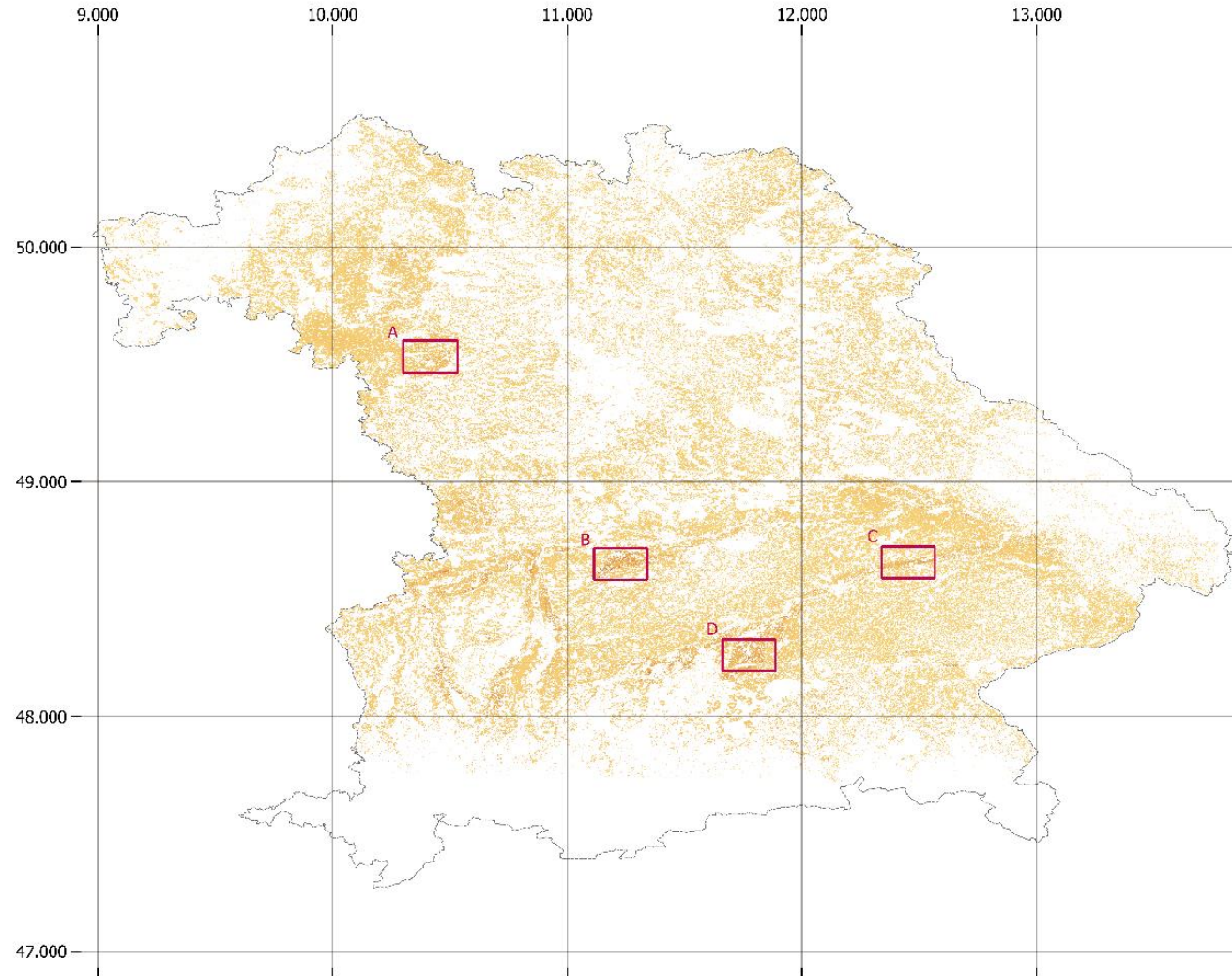
Top and bottom 5% activations per each filter



The **visible range**, and in particular the beginning of the spectrum. This may be attributed to **soil color** and the albedo of the sample which is influenced by the presence of organic matter

The **upper SWIR region**, where absorptions due to the presence of organic materials may also be found.

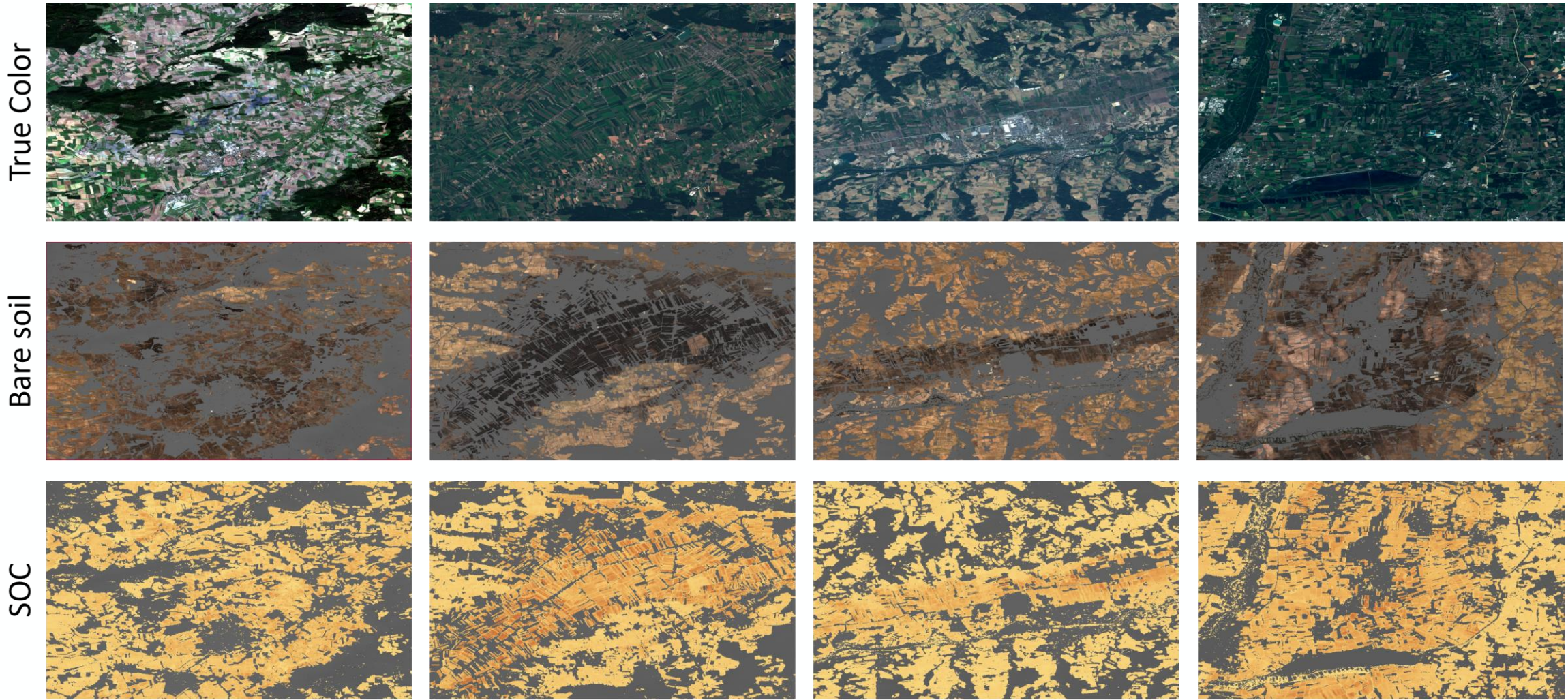
SOC map in Bavaria



20% of the RoI recognized
as exposed soils

Visually homogeneous
predictions; free of any
apparent artifacts

SOC map In Bavaria

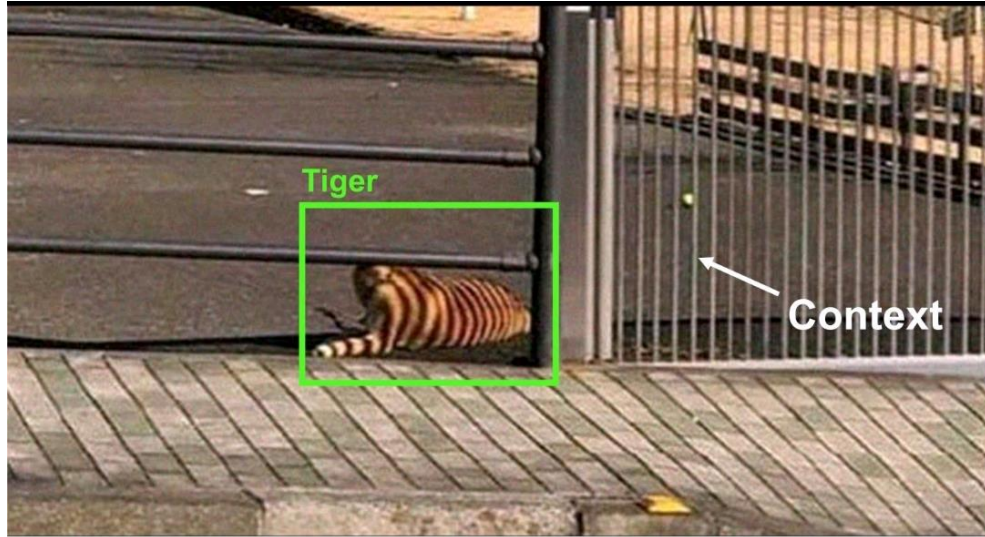


Test site A

Test site B

Test site C

Test site D



Conclusion

We developed for a first time a localized multichannel CNN able to handle Sentinel-2 data to predict soil properties

Future steps

Exploitation of additional information sources like the DEM and environmental covariates in a multi-branch approach;

Utilize both **geographical and spectral distance vectors** for neighbor selection.

Utilize **hyperspectral missions** (e.g., PRISMA, EnMAP) to leverage the benefits of the multi-input CNN model.

THANK YOU!

CONTACT US

ntziolas@auth.gr

Nikolaos Tziolas



ARISTOTLE
UNIVERSITY
OF THESSALONIKI

